FACTORS AFFECTING MALAYSIAN TERTIARY STUDENTS' ENGLISH PROFICIENCY

SHIRLISA YEOH LAY PHING

UTAR

A RESEARCH PROJECT
SUBMITTED IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE BACHELOR OF ARTS (HONS) ENGLISH LANGUAGE
FACULTY OF ARTS & SOCIAL SCIENCES
UNIVERSITI TUNKU ABDUL RAHMAN
APR 2021

FACTORS AFFECTING MALAYSIAN TERTIARY STUDENTS' ENGLISH PROFICIENCY

SHIRLISA YEOH LAY PHING

A RESEARCH PROJECT
SUBMITTED IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE BACHELOR OF ARTS (HONS) ENGLISH LANGUAGE
FACULTY OF ARTS & SOCIAL SCIENCES
UNIVERSITI TUNKU ABDUL RAHMAN
APR 2021

ACKNOWLEDGEMENTS

This research project has been successfully accomplished with the assistance of various parties. It would have been difficult to complete this research project without their assistance, guidance and cooperation. Therefore, I am really grateful to these people for being there for me throughout the whole journey in completing this research project.

I would first like to express my gratitude towards my supervisor, Ms. Cheng Siew May, who gave great guidance and support throughout the duration of this research project. I really appreciate the patience, knowledge, comments and valuable feedback she has given. Her great effort has made it so that this research project was completed smoothly. I would also like to thank my family and friend for their encouragement and understanding. The constant support I get from them is highly valued and appreciated.

To every one of you, I want to express my fullest gratitude. This thesis has been completed by having all of your effort and contributions.

SHIRLISA YEOH LAY PHING

APPROVAL FORM

This research paper attached hereto, entitled “Factors Affecting Malaysian Tertiary Students' English Proficiency” prepared and submitted by “Shirlisa Yeoh Lay Phing” in partial fulfilment of the requirements for the Bachelor of Arts (Hons) English Language is hereby accepted.

_____ Date:_____

Supervisor

Supervisor’s name: Ms. Cheng Siew May

ABSTRACT

The aim of the research was to find out to what extent do code mixing, code switching and first language interference affect university students' English proficiency. For this study, 384 survey questionnaires were collected for the data analysis of the research. From the results obtained by running the gathered data in SPSS Statistics 26, it was found that code mixing, code switching and first language interference do affect university students' English proficiency but to different extents. It was found that code mixing and code switching positively affects university students' English proficiency while first language interference has a negative consequence to university students' English proficiency. The findings also concluded that code switching greatly affects university students' English proficiency while code mixing and first language interference minimally affects university students' English proficiency.

DECLARATION

I declare that the material contained in this paper is the end result of my own work and that due acknowledgement has been given in the bibliography and references to ALL sources be they printed, electronic or personal.

Name: SHIRLISA YEOH LAY PHING

Student ID: 17AAB01913

Signed: 

Date: 13th APRIL 2021

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CM	Code Mixing
CS	Code Switching
DV	Dependent Variable
ECLS-K	Early Childhood Longitudinal Study-Kindergarten Cohort
EP	English Proficiency
IV	Independent Variable
L1	First Language Interference
MLR	Multiple Linear Regression
MUET	Malaysian University English Test
SPSS	Statistical Package for Social Science

CHAPTER ONE
INTRODUCTION

1.0 Introduction

This study aims to investigate to what extent code mixing, code switching, and first language interference affect the English proficiency of university students in Malaysia. This chapter serves as a purpose to understand the foundation of the study by explaining the background of the study, statement of the problem, research objectives, research questions, significance, scope and limitations of the study as well as the definition of terms.

1.1 Background of the Study

English was first introduced in Malaysia when British colonisers made it an important language as well as associating the language with high status and power. However, the language lost its status after Malaysia gained independence in 1957 and made the Malay language its national language. As a result, Malay language is used in government administrations and public education (Pillai & Ong, 2018). Despite that, English still plays a significant role and is taught to students attending national schools from Standard One up to Form Five (Nazri, 2013).

The diversity of races and ethnicities in Malaysia has also contributed to bilingualism or multilingualism. It is natural to find individuals communicating in their native language; however, the language barrier has made it so that a handful Malaysians would use English as a common tool of communication (Song, 2019; Thirusanku & Melor, 2014). On this account, code mixing and code switching is adopted. The use of code mixing and code switching are regarded as a natural habit for those who have other languages as their first language (Kamisah & Misyana, 2011).

1.2 Statement of the Problem

English is a language that is internationally used as a form of communication between individuals of different native language (Thirusanku & Melor, 2012). Even the former Prime Minister, Tun Dr Mahathir Mohamad has acknowledged the importance of the English language and have urged Malaysians to master it (The Star, 2020).

Nonetheless, many Malaysians are still not proficient in English. Thirusanku and Melor (2014) said that although they have studied English for six years in primary education and five years in secondary education, it is hard to believe why some Malaysians are still not comfortable with the language. According to the Malaysian Examinations Council (2017), more than half of the students who sat for Malaysian University English Test (MUET) in November 2016 were unable to obtain band 4 or above. MUET scores are classified from band 1 to band 6, where band 1 denotes "an extremely limited English user" and the range gradually increases to band 6 which denotes "an advance English user" (Zulkifli et al., 2011).

In fact, The Star (2017) also reported that the lack of English language proficiency among Malaysians have been a major cause as to why they do not get job offers from international companies. Employment has become a concern for Malaysian fresh graduates as some graduates are unable to find employment after getting their hard-earned degree. One of the main determinants for this phenomenon is the low level of English proficiency among these graduates.

The low level of English proficiency among Malaysians might be the negative effect of bilingualism or multilingualism in the Malaysian society. In fact, Latisha and Surina (2013) noted that bilingual or multilingual societies commonly code mix or code switch in their daily conversations. These phenomena have left a negative impact on an individual's level of English proficiency (Lau et al., 2011; Sardar et al., 2015).

Additionally, first language interference also plays a vital role in shaping a person's English proficiency level. The vast differences between a person's first language and English are said to be the leading factors as to why an individual failed to excel in the English language (Karim & Nassaji, 2013).

With these problems in mind, this study aims to figure out to what extent do code mixing, code switching, and first language interference affect the English proficiency of university students in Malaysia.

1.3 Research Objectives

The objectives of this research are:

1. to find out to what extent code mixing affects the English proficiency of university students in Malaysia.
2. to find out to what extent code switching affects the English proficiency of university students in Malaysia.
3. to find out to what extent first language interference affects the English proficiency of university students in Malaysia.

1.4 Research Questions

The aim of this research is to answer the following research questions:

1. To what extent does code mixing affect the English proficiency of university students in Malaysia?
2. To what extent does code switching affect the English proficiency of university students in Malaysia?
3. To what extent does first language interference affect the English proficiency of university students in Malaysia?

1.5 Significance of the Study

The results of this study will feature valuable knowledge and understanding on Malaysians' English proficiency as it aims to determine factors that influence Malaysians' English proficiency. The study is able to benefit multiple parties such as the government, educational institutes, teachers and students as the results will pinpoint the factors affecting Malaysians' English proficiency on a large scale.

With this study, the mentioned parties can understand why Malaysians are still lacking in the language and put importance on improving the people's English. They may use this study as a reference to develop better strategies that will enhance Malaysian's English proficiency.

1.6 Scope and Limitations of the Study

The coverage of this study is to analyse the extent to which code mixing, code switching, and first language interference affect the English proficiency of university students in Malaysia. A sample size of 384 respondents will partake in answering a close-ended questionnaire. The gathered data will then be analysed to produce meaningful data that will assist in achieving the objectives of the study.

Some of the limitations identified for this study are time constraint and data collection method. As the researcher was only given a limited time to conduct the study, the researcher will only be using one data collection method. The researcher will be using questionnaires to collect the data needed as it is one of the fastest, easiest and one of the most commonly used tools. However, the inability to utilise various data collection methods may affect the quality and accuracy of the results.

Due to time constraint, the researcher is unable to use other data collection methods. Methods like interviews and focus group discussion will not be used because these methods will require more tedious and time-consuming processes; for instance, looking for

knowledgeable and suitable participants, transcribing and analysing complex data. By using interviews and focus group discussion, the researcher can gather more data and deeper understanding about the participants' experience, opinion and attitude towards the study (Ruth & Marcia, 2018).

1.7 Definition of Terms

According to the University of Southern Queensland (2016), English proficiency is said to be the capability of an individual to utilise the English language as a means of communication. The level of an individual's English proficiency can be measured by taking into account the individual's reading, speaking, writing and listening skills (Geide-Stevenson, 2018).

As for code mixing, it is more popularly defined as the mixture of two different languages in terms of lexical items or grammatical features in a particular sentence (Muysken, 2000). This occurrence can be done in either written or spoken discourse (Iman et al., 2015).

Moreover, code switching is defined as the shift of use between two or more languages in a discourse. This occurrence can happen in a conversation when a person begins their conversation in one language and shifts to another language halfway, or when a person uses one language but receive replies in a different language (Azam & Mahdieh, 2013).

According to Lao (2017), first language interference is the hindrance on a speaker's first language when acquiring the second language. It happens when individuals apply the knowledge of their first language into the second language.

1.8 Conclusion

To sum it up, this chapter introduces the topic of the study by providing the overview of this study. The background of the study, statement of the problem, research objectives,

research questions, significance, scope and limitations of the study and definition of terms had been discussed to provide better understanding on the current situation of the topic as well as justifying the significance of the study. The upcoming chapter will be a literature review on various past studies linked to this study.

CHAPTER TWO

LITERATURE REVIEW

2.0 Introduction

In this chapter, the researcher will provide reviews of past studies that can help readers to further understand the relationship between the dependent variable and the independent variables. Each variable is thoroughly reviewed in order to draw out suitable hypotheses for this research. In addition, The Analysis of Linguistic Borrowing (Haugen, 1950) is also explained as it serves as the foundation of this research.

2.1 English Proficiency

Several researches related to English proficiency were studied. According to Geide-Stevenson (2018), English proficiency of an individual can be measured by evaluating their reading, speaking, writing and listening skills. In her study on foreign students' (who are non-native English speakers) academic performances, she found out that their academic performances are significantly affected by their English proficiency. Those who are not highly proficient in English were unable to perform well probably because they were unable to fully comprehend what they were learning. From this study, it can be deduced that the foreign students' English proficiency were affected by language contact with their native language, which can be classified as first language interference.

On the other hand, Palacios and Kibler (2016) had researched on how languages used at home and in formal education of kindergartens affected students' mastery on oral proficiency and reading mastery of the English language. The study utilised data gathered by Early Childhood Longitudinal Study–Kindergarten Cohort (ECLS-K) on 21,409 kindergarten students in the United States. The researchers first distinguished each student's primary language at home and categorised them as English speakers or non-native English speakers.

They then analysed and further categorised the students' scores from multiple English tests to determine their level of English reading and speaking proficiency among the English speakers and the non-native English speakers. Their study revealed that 45% of the English speakers were proficient in reading. Meanwhile, only 29% of the non-native English speaker were proficient in reading. Moreover, the researchers also discovered that students who are exposed to the English language at a later age had more difficulty in speaking the language fluently. From their study, it can be concluded that English proficiency is affected by a person's first language.

Alternatively, Gilakjani and Reza (2011) studied about the factors influencing English listening comprehension. In their study, they found that English learners have issues when it comes to English listening comprehension. This is because a handful of course books and teachers do not emphasise on listening and speaking skills when coming up with teaching plans. Furthermore, the perception of the importance of listening, the scarce teaching theories on listening and the low use of affective teaching methods for listening are also prime factors influencing English learners' listening comprehension. In another study on the factors influencing English pronunciation, Gilakjani (2012) mentioned attitude, motivation, instruction and exposure as the several prominent factors affecting speakers' English pronunciation.

The three studies conducted by Palacios and Kibler (2016), Gilakjani and Reza (2011) and Gilakjani (2012) had researched on English proficiency. However, those studies had only focused on some of the components used to measure an individual's English proficiency level and failed to research on all components stated by Geide-Stevenson (2018). Hence, for this research, the researcher wishes to include all components used to measure an individual's English proficiency level to obtain more accurate results.

2.2 Code Mixing

Code mixing is a common occurrence in Malaysia due to its nature of being a multilingual society and that has influenced language users when using a language, regardless of whether it is spoken or written. Even Malaysians who are proficient in English are often found to code mix the language with Malay, Chinese or Tamil (Lau et al., 2011).

Iman et al. (2015) proposed that speakers need to have high proficiency in Malay and English in order to code mix. Their study explained that code mixing can only be executed if the speakers have mastered the fundamentals of both languages. This is because speakers will be integrating one language to the other and they will have to do it without distorting the language to the point where it is no longer comprehensible. Having said that, the study also discussed academicians' concern on code mixing. They fear that students might get too used to code mixing and apply it in their formal education.

In contrast to the above study, Kamisah and Misyana (2011) claimed that code mixing occurs due to the lack of English proficiency. In the study, they analysed the frequency and attitudes on code mixing and code switching of English and Malay in a classroom environment. They found that the leading reason for code mixing was the low level of English proficiency. Both students and instructors confessed that they code mix in order to explain or express their opinions when they were unable to do it in English due to their incompetence.

Moving on, Younas et al. (2014) and Rahimi (2014) found that code mixing assists the process of learning the English language and more so when it comes to picking up new vocabularies. In the study conducted by Younas et al. (2014), the researchers discovered that the learners felt more comfortable when their teachers code mix during the teaching process. Their study also shared similar findings with Rahimi (2014), where majority of the learners agreed that it was easier for the them to memorise and understand information when their

teachers code mix. Even though the learners were learning more effectively, both studies pointed out that there was lack of exposure of the English language. Hence, Younas et al. (2014) revealed that learners were unable to fully utilise the writing skills that they acquired, whereas Rahimi (2014) stated that the learners had no problem in recognising the vocabularies they had learned but had difficulty in using them.

2.3 Code Switching

Code switching, like code mixing, is an everyday occurrence among Malaysians who are bilingual or multilingual. Malaysians are found to code switch for numerous reasons such as using it as a medium of communication in the multicultural society and also using it in teaching to help students understand better (Muthusamy et al., 2020).

In different studies where they looked into the relationship between code switching and English language learning, it was revealed that students exhibit positive attitudes if code switching is incorporated during the process of learning the English language. Studies claimed that code switching is a basic need for learners to play an active role in the learning process (Azam & Mahdieh, 2013; Fareed et al.,2016). Azam and Mahdieh (2013) justified that the proper use of code switching in class might generate successful outcomes in language learning.

In spite of that, the study conducted by Fareed et al. (2016) found abundant of disadvantages of code switching. One of the main disadvantages discussed in the study was the lack of exposure to the English language. Some teachers knew no bound and are found to heavily code switch in classes. Those events had led to a series of problems. First, the lack of exposure to the English language had negatively affected learners' fluency in speaking English. It was also found that they had poor listening skills. Learners struggled to understand and fully comprehend English when there was an increase in speed. Next, learners were unable to pick up new vocabularies and this would greatly affect the process of becoming proficient in a

language. Therefore, the lack of exposure to English was an obstruction to learners who wanted to improve their speaking skill, listening skill and writing skill (Fareed et al., 2016; Tsukamoto, 2011).

Apart from that, a study by Lee et al. (2013) stated that the practice of code switching in class was negatively perceived. Their study was carried out on 27 software engineering students to analyse whether code switching contributes to the success of learning. They discovered that code switching does not assist students' learning process. Also, the students preferred their classes to be conducted fully in English and hoped that their lecturers would not code switch.

In another study conducted by Sardar et al. (2015) on Iraqi students attending a university in Malaysia, they claimed that frequent use of code switching will negatively influence an individual's English proficiency. This is because code switching acts as an obstacle for them to polish their English. The study also uncovers the reasons behind the Iraqi students' use of code switching. Among them were their low level of English proficiency and lack of confidence in speaking English.

2.4 First Language Interference

In Malaysia, most Malaysians use either the national language (Malay language), their mother tongue or the main dialect spoken in their area as their first language while English is taught as a second language to all students from Standard One up to Form Five as a compulsory subject (Song, 2019; Nazri, 2013). This is because English plays a vital role and it is used for many purposes such as cross-cultural communications, education in tertiary studies and works in many national and international firms (Nazri, 2013).

There are studies that claimed that first language interference facilitates the acquisition of the English language which leads to better English proficiency (Ochi, 2009; Madriñan, 2014).

Ochi (2009) claimed that the use of first language via Interpreting Training Methods has brought positive output to English learners' performance. He carried out a test on 33 Japanese and the generated results proved that task comprehension in Japanese (their first language) facilitated better English performance by lowering students' anxiety level when learning a different language. Madriñan (2014) conducted a study to find out whether the use of first language in class would assist in the English language acquisition process. The study confirmed that students can take advantage of their first language when learning English by applying similar concepts from the first language to the English language.

However, vast studies have affirmed that there is a significant relationship between first language interference and English proficiency. There are a few previous studies that showed that first language interference negatively affects an individual's English proficiency (Sinha et al., 2009; Lao, 2017). According to Sinha et al. (2009), speakers' fluency in the English language among Asians like Chinese, Indians and Koreans are lower due to the shapes and structural differences in alphabets between their first language and English. In Lao's (2017) study, he conducted a research on how first language affects the English writing skills of students of age 12 to 13 in a school. He found that the first language affects the students' writing skill in four different areas namely morphology, phonology, semantics and syntax.

2.5 The Analysis of Linguistic Borrowing (Haugen, 1950)

This research will be carried out using Haugen's Analysis of Linguistic Borrowing (1950), often quoted as the model of lexical borrowing as its foundation. This is because Malaysian English (Manglish) is heavily influenced by the Malaysian's first language and incorporates code mixing and code switching in its usage. This happening affects the level of English proficiency among Malaysians (Thirusanku & Melor, 2013).

Haugen's (1950) work signalled the start of the modern interests on the studies of lexical borrowing among bilinguals (Hoffer, 2002, as cited in Mohammed, 2019). His work is known to be significant in studies with bilingualism, language contact and also lexical borrowing (Bahumaid, 2015).

Andersen et al. (2017) stated that lexical borrowing is the root of code mixing and code switching. Although lexical borrowing begins as single word switches, it progressively becomes a normal norm that could lead to code mixing and switching (Haspelmath & Tadmor, 2009). The diversity of ethnicity in Malaysia creates a multilingual society which makes it possible for Malaysians to code mix or code switch (Vollmann & Soon, 2019). In Haugen's (1950) study, he explained borrowing as a process that occurs when a bilingual speaker adopts morphemes, pronunciation or structure from one language to be used in another language.

Haugen (1950) first explained the two processes of borrowing which are importation and substitution. Importation is said to be the borrowing of words from one language to another without changing the natural form of the words. As for substitution, words from one language undergo some changes in their natural forms before they are borrowed to another language (Malah, 2014).

Haugen (1950) then classified three different types of word borrowings namely, loanwords, loanblends and loanshifts. Loanwords are importation; hence, the words do not undergo any changes and are borrowed as it is by other languages (Malah, 2014). From Tan's (2009) study, it can be observed that English in Malaysia have many loanwords such as *congkak, syok* and *kenduri.* Loanblends, on the other hand, are importation and substitution where they are made up of a morpheme or word from both languages involved, such as *batik cloth* and *ter-sleep* (Malah, 2014; Tan, 2009). As for loanshifts, they are substitution where there are changes in meaning, pronunciation or grammatical structure in the words (Haugen,

1950). In Manglish context for instance, the word *gostan* was derived from the phrase *go astern* which carries the meaning *reserve a vehicle* (Tan, 2009).

2.6 Hypotheses Development

Hypothesis 1:

H_0: English proficiency has no significant relationship with code mixing.

H_1: English proficiency has a significant relationship with code mixing.

The relationship between the English proficiency and code mixing among university students in Malaysia is expected to be a significant negative one. Although code mixing will assist the learning process, students were unable to fully utilise what they had learned in class, which means that their English proficiency level is affected (Younas et al., 2014; Rahimi, 2014).

Hypothesis 2:

H_0: English proficiency has no significant relationship with code switching.

H_1: English proficiency has a significant relationship with code switching.

According to Fareed et al. (2016) and Sardar et al. (2015), constant use of code switching decreases the use of English, which will then affect speakers' English proficiency. Hence, the relationship between English proficiency and code switching is expected to be a significant negative one among university students in Malaysia.

Hypothesis 3:

H_0: English proficiency has no significant relationship with first language interference.

H_1: English proficiency has a significant relationship with first language interference.

A significant negative relationship is expected to exist between the English proficiency and first language interference among university students in Malaysia. In line with studies by

Sinha et al. (2009) and Lao (2017), the researcher believes that the influence of an individual's first language will negatively affect their English proficiency.

2.7 Conclusion

In brief, this chapter presents a review on past literatures in order to better understand the relationship between English proficiency and the factors influencing it namely code mixing, code switching and first language interference. Furthermore, a relevant theoretical model relevant to this study was provided and hypotheses were developed for each of the independent variable. In the next chapter, the researcher will discuss the methods that will be employed to conduct this research.

CHAPTER THREE

METHODOLOGY

3.0 Introduction

This chapter discusses on the methods undertaken by the researcher to figure out to what extent code mixing, code switching, and first language interference affect the English proficiency of university students in Malaysia. The researcher also explains the reasons as to why certain tests will be carried out for data analysing.

3.1 Research Design

This research is a quantitative analysis that will be employing primary data collection by using close-ended questionnaires as the research instrument. This method has been chosen because it can gather a large amount of information from the targeted population in a short period of time. The data collected will then be quantified in order to create a meaningful data. This will allow the researcher to analyse and determine to what extent do code mixing, code switching, and first language interference affect the English proficiency of university students in Malaysia.

3.2 Population and Sample

The target population for the research will be university students in Malaysia. This research will utilise random sampling as the sampling technique. The sample size was determined by using Krejcie and Morgan (1970) table of determining sample size for research activities as shown below.

Table 3.1

Krejcie and Morgan (1970) table of determining sample size for research activities

Population Size	Sample Size at 95% Confidence Level and 5% Margin of Error
50,000	381
100, 000	383
500, 000	384
1, 000, 000	384
2, 500, 000	385

Source. From The Research Advisors, (2006).

According to New Straits Times (2019), it is estimated that there are approximately 1.3 million youths who are enrolled in tertiary education. As there is a population of 1.3 million, the research will need a sample size of 384 university students, which gives 95% level of confidence with a margin error of 5%. A pilot test will be conducted before the actual research is carried out to test the accuracy of the questions in the questionnaires and make any necessary amendments. This will avoid unwanted errors during the actual research.

3.3 Data Collection

To gain more accurate results, both primary and secondary data will be used in this study. Closed-ended questionnaires will be used for data collection because they are more effective and efficient in gathering data from a large sum of respondents. The collection of data will be obtained through distributing the questionnaires to university students in Malaysia via Google Form. The secondary sources that will be used are websites, journals and articles. Different online sources will be utilised to gather relevant information that will assist the completion of the research.

In Part 1 of the questionnaire, the researcher will ask questions regarding the respondents' demographic profile to understand the characteristics of the respondents. This will allow the researcher to characterise the respondents and analyse the data gathered. Part 2 of the questionnaire will contain questions to get results on the respondents' level of English proficiency. Both Part 1 and Part 2 of the questionnaire will be adapted from Ahmad and Shaima (2016). Lastly, Part 3 of the questionnaire will relate the respondents' English proficiency to the factors influencing them. This part will be covering the independent variables – code mixing, code switching and first language interference. Ahmad and Shaima's (2016) questionnaire will be adapted especially on questions related to code mixing and code switching, while questions relating to first language interference will be adapted from Alia et al. (2019). The questionnaire will utilise several measurement scales including ordinal scale, nominal scale and five-point Likert scale.

3.4 Data Analysis

In order to produce a more meaningful information, data analysis will be carried out using the data collected. The responses gathered from the respondents will be checked thoroughly to ensure errors can be detected on earlier stage. Any errors identified will either be adjusted or omitted to maintain the quality of the data.

Then, the data will be coded by grouping the responses into several classifications before assigning numbers to them. As an example, the respondents' gender will be coded as '1' for male and '2' for female. Later, Statistical Package for Social Science (SPSS) software will be used to analyse the coded data. By using this software to analyse the data, it enables data management and statistics calculation based on the data collected from the questionnaires.

After that, reliability test will be carried out to test the legibility of the data. Cronbach's alpha internal consistency will be implemented to measure the scale of reliability. The rule of thumb for interpreting Cronbach's alpha internal consistency is shown in Table 3.2 below.

Table 3.2

The rule of thumb for interpreting Cronbach's alpha internal consistency

Cronbach's alpha	Internal consistency
$\alpha \geq 0.9$	Excellent
$0.9 > \alpha \geq 0.8$	Good
$0.8 > \alpha \geq 0.7$	Acceptable
$0.7 > \alpha \geq 0.6$	Questionable
$0.6 > \alpha \geq 0.5$	Poor
$0.5 > \alpha$	Unacceptable

Source. From Chaudhary, (2016).

Next, a normality test called the Shapiro-Wilk test is carried out to test the normality of the data to fulfil the need of a normal distribution. The statistic value must be more than 0.05 to ensure that the data is normally distributed (Razali &Wah, 2011).

Furthermore, a correlation analysis will be carried out to measure the linear relationship of the two variables. There is a range of -1 to +1 in Pearson's correlation coefficient. When there is a positive sign, there will be a positive correlation, and when there is a negative sign, there will be a negative correlation. Subsequently, 0 indicates no correlation. Additionally, the way to interpret the Pearson's correlation coefficient is when the coefficient ranges from 0.1 to 0.3 or from -0.1 to -0.3, the correlation is considered weak. Meanwhile when the coefficient

ranges from 0.3 to 0.5 or from -0.3 to -0.5, the correlation is considered medium and when the coefficient is more than 0.5 or less than -0.5, the correlation is considered strong (Sedgwick, 2012).

Moving on, multiple linear regression (MLR) model will be used to explain the association between a continuous dependent variable and its independent variables (Smalheiser, 2017). R-square is used to determine an appropriate linear regression model by assessing the association between the English proficiency of university students in Malaysia and the model within the range of 0% to 100% (Frost, 2017).

The equation form of the multiple linear regression is shown below:

$$Y=\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots\cdots+\beta_n X_n \qquad (1)$$

Multiple linear regression will be used to investigate whether the three independent variables significantly affect the English proficiency of university students in Malaysia. Thus, the equation form of this study is:

$$EP=\beta_0+\beta_1 CM+\beta_2 CS+\beta_3 L1 \qquad (2)$$

whereby,

EP = English proficiency of university students in Malaysia (Dependent Variable)

CM = Code mixing (Independent Variable 1)

CS = Code switching (Independent Variable 2)

L1 = First language interference (Independent Variable 3)

3.5 Conclusion

To conclude, this chapter describes the procedures planned out by the researcher. This research will gather needed data by means of close-ended questionnaires that will be distributed to 384 university students in Malaysia. The gathered data will then be run using SPSS.

Therefore, this chapter serves as a clear guideline that will assist the researcher throughout the duration of the research. The next chapter will display the output generated along with a detailed analysis of the findings of the research.

CHAPTER FOUR

FINDINGS & ANALYSIS

4.0 Introduction

In this chapter, the results of the research are stated and analysed. The results of both the pilot tests and the actual results of this research are first generated by SPSS and later analysed by the researcher. Results included in this chapter are respondents' demographic profile, central tendency and standard deviation, reliability, normality, Pearson's correlation coefficient and the multiple linear regression.

4.1 Research Questions

The analysis of data on the gathered data were carried out to address the following research questions below:

1. To what extent does code mixing affect the English proficiency of university students in Malaysia?
2. To what extent does code switching affect the English proficiency of university students in Malaysia?
3. To what extent does first language interference affect the English proficiency of university students in Malaysia?

4.2 Pilot Tests

To carry out the pilot tests, 30 sets of data were collected. The data were gathered, and the results were run using SPSS before the researcher further proceeded with the study. All relevant tests were carried out to ensure the validity and reliability of the study. The results of the pilot testes are stated in Table 4.1 and 4.2.

Table 4.1

Summary of the research's pilot tests

Tests	Results				
	All Variables	EP	CM	CS	L1
Reliability test	0.725	-	-	-	-
Normality test (p-value)	-	0.058	0.067	0.112	0.066
R-square	0.233	-	-	-	-
Adjusted R-square	0.145	-	-	-	-
F-test (p-value)	0.071	-	-	-	-
t-test (p-value)	-	0.000	0.038	0.016	0.678
Parameter estimates	0.413 (Constant)	-	0.194	0.219	0.117

Note. Developed for research.

Table 4.2

Pilot test for Pearson's correlation coefficient

	EP	CM	CS	L1
EP	1.000	0.669**	-0.595**	-0.174
CM		1.000	0.746**	0.504**
CS			1.000	0.573**
L1				1.000

**. Correlation is significant at the 0.05 level (2-tailed).

Note. Developed for research.

Based on Table 4.1, the Cronbach's alpha value is 0.725, which means that the reliability of all the 30 data entries is considered to be acceptable as it is valid and reliable. Besides, all the p-values for the normality tests were more than 0.05. This means that the data is normally distributed.

On the other hand, the Pearson's correlation coefficient shown in Table 4.2 indicates that all the variables are correlated to one another and have a strong correlation, with the exception of the correlation between English proficiency and first language interference, which displays a weak correlation. In addition, all the Pearson's correlation coefficient values were less than 0.9, hence, there is no multicollinearity issue on these sets of data.

In accordance with Table 4.1, the R-square value is 0.233. This means that 23.3% of the level of English proficiency among university students in Malaysia can be explained by code mixing, code switching and first language interference. However, the p-value of the F-test is more than 0.05 which means that the model for this 30 sets of data is not fit.

Moving on, the p-values for the t-test for all the variables except for first language interference are less than 0.05. In other words, it is shown that apart from first language interference, both code mixing and code switching does affect the level of English proficiency among university students in Malaysia. The equation form is shown as:

$$EP = 0.413 + 0.194\,CM + 0.219\,CS \qquad (3)$$

Equation 3 above represents the relationship between the independent variables and the dependent variable for the pilot test. According to the pilot test for this study, the equation can be explained by with every increase in one unit of code mixing and code switching, it will affect the English proficiency of university students in Malaysia to increase by 19.40% and 21.90% respectively, holding other variables constant. In other words, this means that code

mixing positively affects the level of English proficiency by 19.4% while code switching positively affects the level of English proficiency by 21.9%. In contrast, first language interference is insignificant to the level of English proficiency of university students in Malaysia.

4.3 Respondents' Demographic Profile

The demographic profiles of the survey questionnaires' respondents are summarised in this section. The frequency analysis was performed by utilising 384 sets of complete and effective questionnaires.

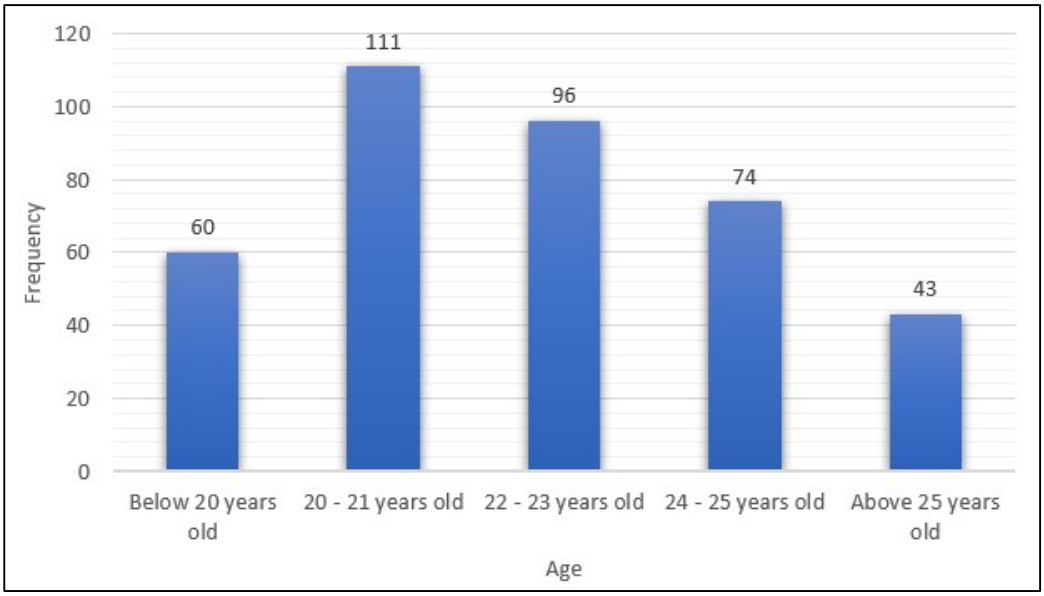

Figure 4.1. Frequency analysis for age.

Note. Developed for research.

The respondents' age are illustrated in Figure 4.1, among which 60 are below 20 years old, 111 are between 20 to 21 years old, 96 are between 22 to 23 years old, 74 are between 24 to 25 years old while 43 are above 25 years old.

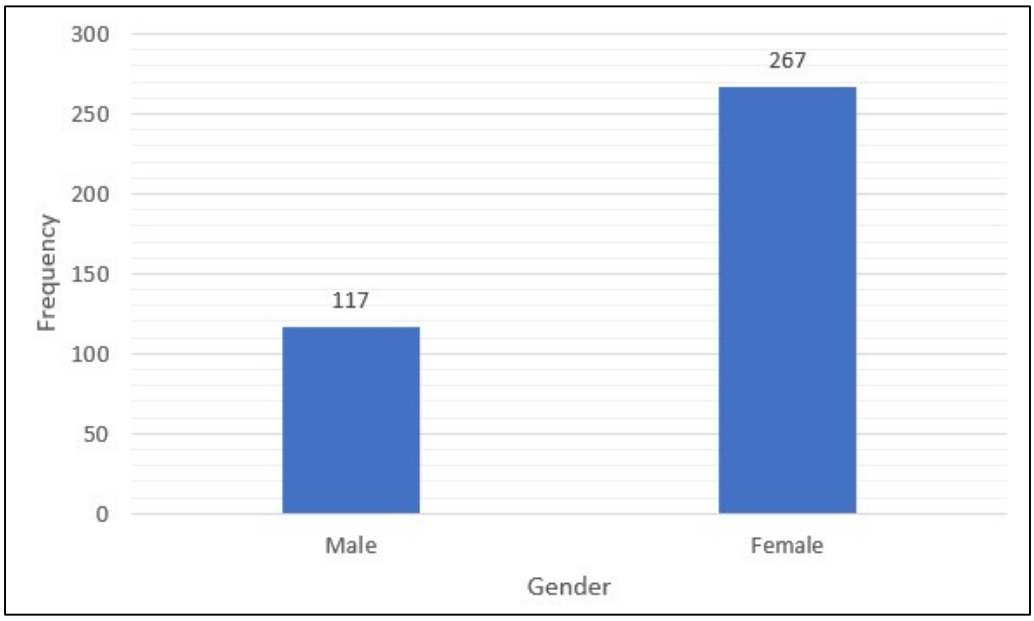

Figure 4.2. Frequency analysis for gender.

Note. Developed for research.

From Figure 4.2, it is clearly shown that the number of female respondents in the survey questionnaires outnumbered the number of male respondents. The number of female respondents stands at 267 while the number of male respondents stands at 117.

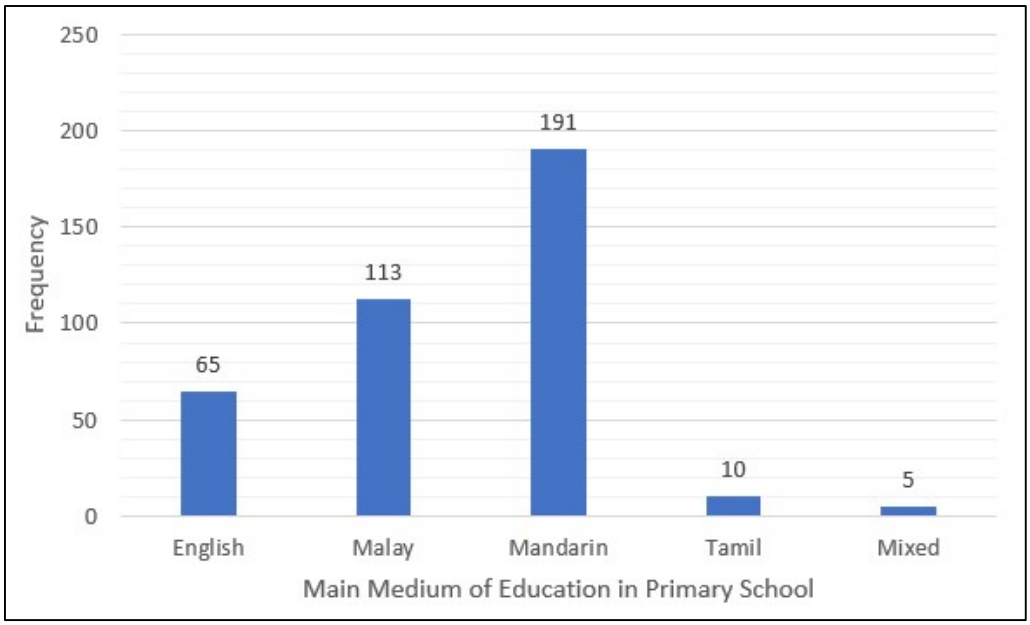

Figure 4.3. Frequency analysis for main medium of education in primary school.

Note. Developed for research.

Figure 4.3 displays the respondents' main medium of education in primary school. Most respondents came from Mandarin medium schools, which accounted for 191 respondents, while 113 respondents came from English medium schools. Additionally, 65 respondents came from Malay medium schools whereas 10 came from Tamil medium schools. On the other hand, 5 respondents came from schools which uses two or more languages as its main medium of education.

4.4 Measures of Central Tendency and Standard Deviation of Constructs

Central tendency measures the mean, median and mode of a data and standard deviation measures the distribution of a set of data (Bhandari, 2020). It is important in helping the researcher to understand the data when descriptive statistics is performed. The central tendency and standard deviation of the 384 sets of data are presented in Table 4.3.

Table 4.3

The measures of central tendency and standard deviation of constructs

Construct	Mean	Median	Mode	Standard deviation
EP	2.9495	3.0000	3.00	0.85896
CM	2.4818	2.5000	2.50	0.42058
CS	2.4870	2.5000	2.50	0.46429
L1	2.4779	2.5000	2.50	0.39724

Note. Developed for research.

According to Table 4.3, the variables have mean scores within the range of 2.47 to 2.95. First language interference has the lowest mean score with 2.4779 while English proficiency has the highest mean score of 2.9495. On the contrary, all the independent variables have a median and mode of 2.50 while the dependent variable has a median and mode of 3.00.

Moreover, for the spread of data in standard deviation, the values for all the variables range from 0.39 to 0.86. English proficiency has the largest spread of data among the variables with its value of standard deviation being 0.85896. In contrary, first language interference has the lowest standard deviation of 0.39724 which means it has the smallest spread of data among all the variables.

4.5 Reliability Test

A reliability test was carried out to test the legibility and validity of the data collected. Cronbach's alpha of internal consistency was implemented to measure the scale of reliability. According to Chaudhary (2016), the Cronbach's alpha must at least be at 0.7 for the data to be considered reliable.

Table 4.4

Reliability test

Cronbach's alpha	Number of items
0.806	3

Note. Developed for research.

Based on Table 4.4, the generated result for the reliability test is 0.806, which is greater than 0.7. According to the rule of thumb for internal consistency, the legibility and validity of

the sets of data collected for this research are good (Chaudhary, 2016). Hence, the questions in the survey questionnaire for this research are reliable and the output generated can be trusted.

4.6 Normality Test

The Shapiro-Wilk test was conducted to test the normality of the data. This is to make sure that the data gathered are normally distributed. The probability value (p-value) must be more than 0.05 to ensure that the data is normally distributed (Razali &Wah, 2011).

Table 4.5

Normality test

Variables	Shapiro-Wilk (p-value)
EP	0.000
CM	0.000
CS	0.000
L1	0.000

Note. Developed for research.

With regards to the output shown in Table 4.5, none of the p-values are more than 0.05. Although the Shapiro-Wilk test is the most commonly used test, when using SPSS and a large sample size is utilised (more than 50 samples), the results generated are inaccurate due to the sensitivity of the test (Elliott & Woodward, 2007). According to Ghasemi & Zahediasl (2012), it is recommended that the normal distribution of a data should be assessed through visual methods and statistical tests to ensure the validity of the normality test. Hence, the researcher applied quantile-quantile plot (Q-Q plot) as a visual method to assist in determining the

normality of the gathered data as the sample size is 384 which is significantly large. The Q-Q plots of the variables are shown below.

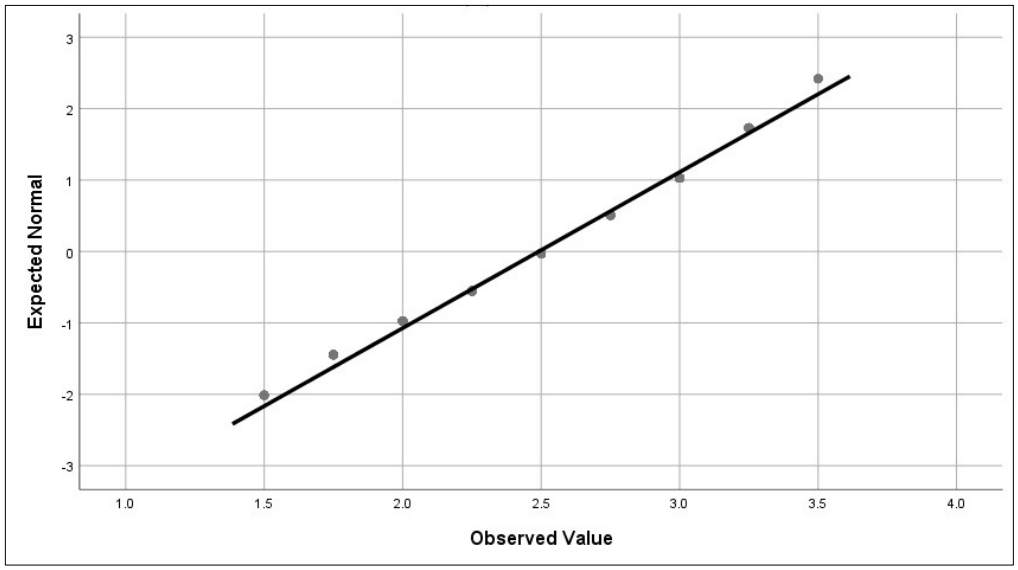

Figure 4.4. Normal Q-Q plot of English proficiency.

Note. Developed for research.

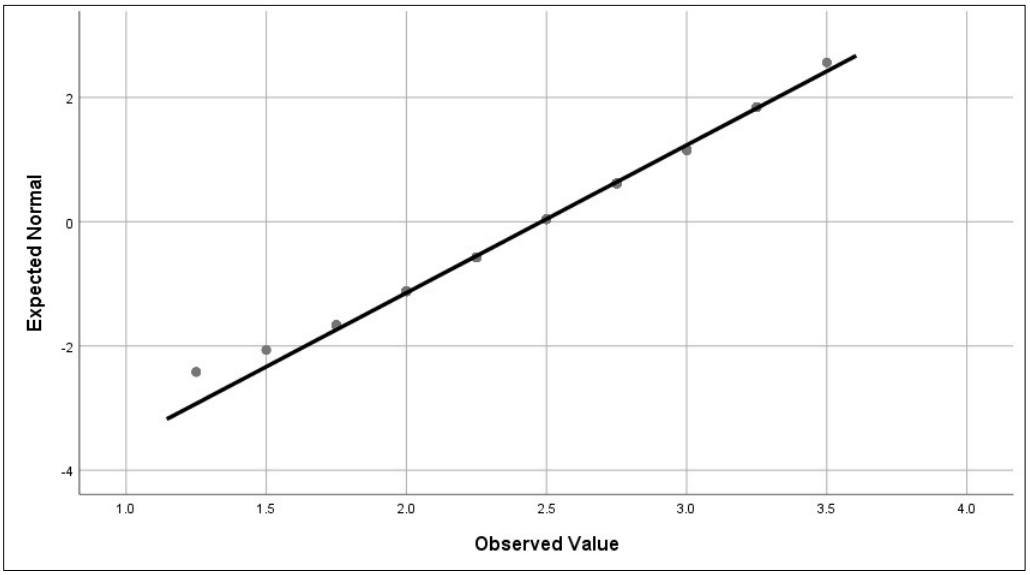

Figure 4.5. Normal Q-Q plot of code mixing.

Note. Developed for research.

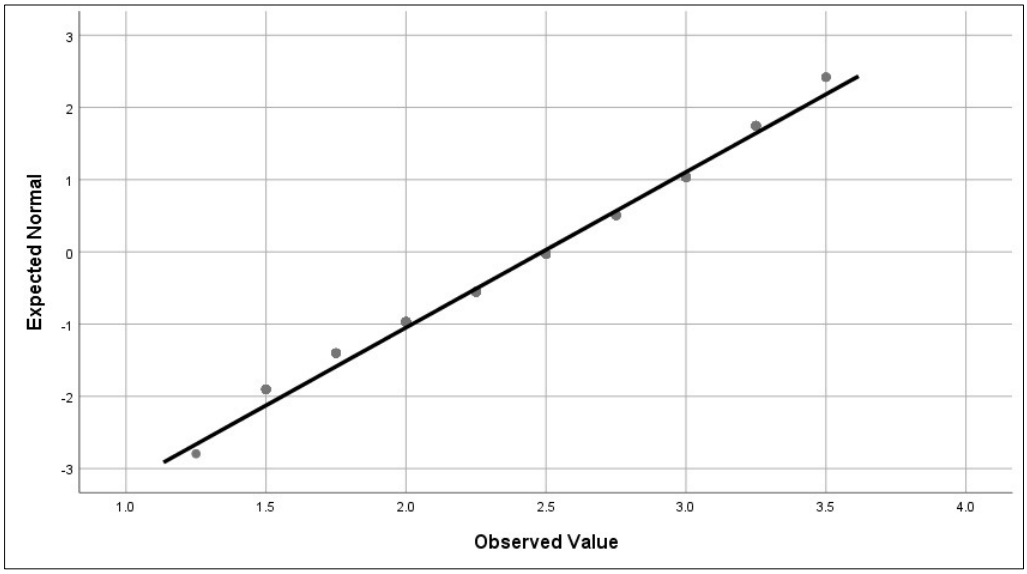

Figure 4.6. Normal Q-Q plot of code switching.

Note. Developed for research.

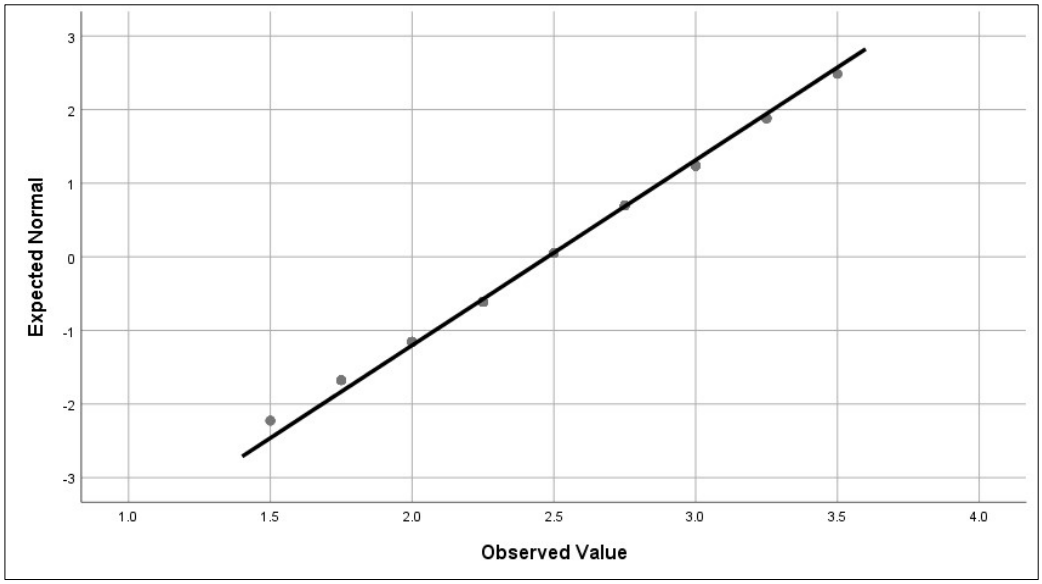

Figure 4.7. Normal Q-Q plot of first language interference.

Note. Developed for research.

From Figures 4.4, 4.5, 4,6 and 4.7, it can be seen that all the variables are actually of normal distribution. This is because most of the points plotted on the graphs are either lying on

the line or deviate slightly from it. The slight deviations are negligible which makes the normal distribution of the data perfectly fit (Varshney, 2020).

4.7 Pearson's Correlation Coefficient

Pearson's correlation coefficient was carried out to test the relationship between all the variables. Additionally, this test can be used to ensure that there is no multicollinearity issue in the data (Sedgwick, 2012).

Table 4.6

Pearson's correlation coefficient

	EP	CM	CS	L1
EP	1.000	0.583**	0.895*	0.670**
CM		1.000	0.575**	0.489**
CS			1.000	0.679**
L1				1.000

**. Correlation is significant at the 0.05 level (2-tailed).

Note. Developed for research.

In accordance with Table 4.6, the Pearson's correlation coefficient indicates that independent variables code mixing, code switching and first language interference have a strong correlation with English proficiency. There is also a strong correlation between code switching and first language inference. In addition, the correlation between code mixing and code switching is strong while the correlation between code mixing and first language inference is medium.

Nevertheless, all the Pearson's correlation coefficient values are significant which means that an increase or decrease of each variable will affect other variables. Moreover, as all

the Pearson's correlation coefficients were less than 0.9, it shows that there is no multicollinearity issue in this research.

4.8 Multiple Linear Regression (MLR)

Multiple linear regression model was used to explain the relationship between a continuous dependent variable (English proficiency of university students in Malaysia) and its independent variables (code mixing, code switching and first language interference) as well as to establish a linear equation for this research (Smalheiser, 2017).

Table 4.7

Model summary

R Square	Adjusted R Square	F-test	Probability of F-test (p-value)
0.991	0.991	13382.843	0.000

Note. Developed for research.

According to Table 4.7, the R Square value which is 0.991 signifies that 99.1% of the changes in English proficiency of university students in Malaysia can be explained by code mixing, code switching and first language interference. Additionally, the probability of F-test (p-value) is below 0.05 and the value of F-test is very large. This indicates that the model in this research is fit.

Table 4.8

Coefficients

Variables	Parameter estimate	t-test	Probability of t-test (p-value)
(Constant)	0.043	2.637	0.009
CM	0.021	3.067	0.002
CS	0.980	135.916	0.000
L1	-0.016	-1.978	0.049

Note. Developed for research.

Table 4.8 displays that all the probability of t-test (p-value) is below 0.05. Thus, code mixing, code switching and first language interference significantly affects the English proficiency of university students in Malaysia.

According to the parameter estimates in Table 4.7, the equation formed for the linear equation is shown as:

$$EP = 0.043 + 0.021\ CM + 0.980\ CS - 0.016\ L1 \quad (4)$$

The equation above represents the relationship between the dependent variable and the independent variables for this research. The positive relationship between English proficiency of university students in Malaysia and code mixing and code switching as well as the negative relationship between English proficiency of university students in Malaysia and first language interference are indicated by Equation 4.

The result can be explained by with every increase in one unit of code mixing, code switching and first language interference, the English proficiency of university students in Malaysia will increase by 2.1%, 98.0% and decrease by 1.6% respectively, holding other variables constant. In other words, this means that code mixing positively affects the level of

English proficiency of university students in Malaysia by 2.1% while code switching positively affects the level of English proficiency of university students in Malaysia by 98.0%. In contrast, first language interference negatively affects the level of English proficiency of university students in Malaysia by 1.6%.

4.9 Conclusion

Chapter 4 displayed all the results that were significant to complete this research. Furthermore, all the results were explained, and a regression was established to address all the research questions. Besides, all hypotheses generated in Chapter 2 are acknowledged and will be discussed in the next chapter. Therefore, Chapter 5 will further discuss on the analysis and findings of the research.

CHAPTER FIVE

DISCUSSION & CONCLUSION

5.0 Introduction

This chapter gives a comprehensive outline from the results obtained as well as discussing the findings of this research with reference to Chapter 4. Furthermore, this chapter presents the research limitations and provide recommendations for any upcoming studies on this topic. Finally, a complete conclusion for this research is made.

5.1 Summary of Statistical Analysis

For this study, the researcher has utilised 384 sets of questionnaires that were complete and effective. Based on the results generated by SPSS, a summary of the multiple linear regression analysis can be seen below in Table 5.1.

Table 5.1

Summary of MLR analysis

Hypothesis	Test IV with DV	Probability of t-test (p-value)	Significance of correlation	Result	Degree of Effect of IV on DV
Hypothesis 1	English proficiency has a significant relationship with code mixing.	0.000 < 0.050	Significant	Reject H_0	2.1%
Hypothesis 2	English proficiency has a significant	0.002 < 0.050	Significant	Reject H_0	98.0%

	relationship with code switching.				
Hypothesis 3	English proficiency has a significant relationship with first language interference.	0.049 < 0.050	Significant	Reject H_0	-1.6%

Note. Developed for research.

From Table 5.1, it can be seen that the independent variables code mixing, code switching and first language interference significantly correlates with the English proficiency of university students in Malaysia as the p-values for these variables are lesser than 0.05. Therefore, H_0 is rejected for hypotheses 1, 2 and 3 as English proficiency has a significant relationship with code mixing, code switching and first language interference.

5.2 Discussion of Findings

Research Question 1: To what extent does code mixing affect the English proficiency of university students in Malaysia?

Hypothesis 1:

H_1: English proficiency has a significant relationship with code mixing.

To address the first research question with regards to Hypothesis 1, it can be concluded that English proficiency has a positive significant relationship with code mixing. According to Table 5.1, the p-value for code mixing is below 0.05. Thus, in this research, the null hypotheses (H_0) is rejected. Moving on, the results are also in contrast with the researcher's expectation in the beginning of the research as code mixing is said to improve Malaysian university students'

English proficiency by 2.1%. This result corresponds with a few past studies mentioned in Chapter 2 (Younas et al., 2014; Rahimi, 2014).

As claimed by Younas et al. (2014) and Rahimi (2014), code mixing assists the process of learning the English language. Learners finds it easier to pick up new vocabularies, memorise and understand the language due to code mixing. Nonetheless, learners' lack of exposure to the English language does hinder the learning process especially in writing and speaking. Hence, code mixing only positively affects the English proficiency of university students in Malaysia minimally at 2.1%.

Research Question 2: To what extent does code switching affect the English proficiency of university students in Malaysia?

Hypothesis 2:

H_1: English proficiency has a significant relationship with code switching.

To address the second research question with regards to Hypothesis 2, it can be concluded that English proficiency has a positive significant relationship with code switching. Table 5.1 shows that the p-value for code switching is below 0.05. Hence, the null hypotheses (H_0) is rejected. The results generated also contradict with the researcher's expectation in the beginning of the research as it can be seen that code switching improves Malaysian university students' English proficiency by 98.0%. The result is in line with one of the previous studies stated in Chapter 2 (Azam & Mahdieh, 2013).

As claimed by Azam and Mahdieh (2013), the proper use of code switching in class might help in generating a successful learning outcome as learners exhibit more positive attitudes. Besides that, code switching enables learners to play an active role which greatly aids the learning process as they are able to use the language when picking it up. Therefore, code

switching positively affects the English proficiency of university students in Malaysia greatly by 98.0%.

Research Question 3: To what extent does first language interference affect the English proficiency of university students in Malaysia?

Hypothesis 3:

H_1: English proficiency has a significant relationship with first language interference.

To address the third research question with regards to Hypothesis 3, it can be concluded that English proficiency has a negative significant relationship with first language interference. Based on Table 5.1, the p-value for code mixing is below 0.05. Hence, the null hypotheses (H_0) is rejected. In addition, the results obtained are according to the researcher's expectation in the beginning of the research as first language interference is said to negatively affect Malaysian university students' English proficiency by 1.6%. This result is consistent with a few past studies mentioned in Chapter 2 (Sinha et al., 2009; Lao, 2017).

As claimed by proficiency Sinha et al. (2009) and Lao (2017), proficiency in the English language among Asians are lower due to the differences between their first language and English. In addition, first language affects students' skill in four different areas namely listening, speaking, writing and reading. However, as Malaysian university students have been exposed to the English language since primary school, their first language does not significantly hinder their English proficiency. Thus, first language interference has a minimal negative effect on the English proficiency of university students in Malaysia with 1.6%.

After conducting the study, it was found that code switching greatly affected the English proficiency of university students in Malaysia, while code mixing and first language interference only affected the English proficiency level minimally. Additionally, it was discovered that code mixing and code switching have positive impacts on an individual's

English proficiency. However, first language interference negatively affects the English proficiency level. In short, code mixing, code switching and first language interference can affect a person's level of English proficiency.

5.3 Limitations and Recommendations

Several limitations that were identified when conducting this research are time constraint and data collection method. As the researcher was only given a limited time to conduct the study, the research could not be conducted on a bigger population such as a country. This is because a bigger population will require a bigger sample size and will need more time to collect the required data. Furthermore, the researcher utilised only one data collection method, which is questionnaire, to collect the data needed because it is one of the fastest, easiest and the most used tool when facing time limitation. Nevertheless, the inability to utilise various data collection methods may affect the quality and accuracy of the results.

There are several aspects which can be considered to further improve the research. In future researches, the researchers can use other data collection methods such as interviews and focus group discussions. By using interviews and focus group discussions, the researchers would be able to gather more data and have a deeper understanding on the participants' experiences, opinions and attitudes towards the topic of research (Ruth & Marcia, 2018). Hence, the researchers would have various types of data which would increase the accuracy and reliability of the data.

Besides, instead of collecting questionnaires to identify the respondents' level of English proficiency, the researcher could conduct English proficiency tests themselves. The researchers will get to personally evaluate the participants' English proficiency. This allows a more standardised evaluation on the participants' proficiency rather than letting the participants

evaluate their own proficiency as some participants might overestimate or underestimate their English proficiency level.

5.4 Conclusion

The results of the research are functional and informative to the education industry as they can take advantage of the results to come up with new policies or strategies to improve Malaysians' English proficiency level. Through the results, government can make improvements and adjustments to existing education policies by taking into consideration of the factors affecting Malaysians' English proficiency level.

Upon completing this study, certain flaws were identified, and recommendations were provided for the deficiencies. These recommendations may assist future researchers who are interested to further look into this topic. Besides, this topic can be expanded to conduct further analyses to develop plans on how to improve Malaysians' English proficiency through code mixing, code switching and first language interference.

REFERENCES

Ahmad, M. A., & Shaima, Q. (2016). A sociolinguistic survey on code switching & code mixing by the native speakers of Bangladesh. *Journal of Manarat International University, 6*(1), 1-17.

Alia, N. A., Melor, M. Y., Harwati, H., Siti, S. S., Nursyuhada, A. S., Zainie, I., & Alia, N. A. Z. (2019). Interference of first language in secondary school students. *International Journal of Engineering and Advanced Technology, 8*(6S3), 675-681.

Andersen, G., Furiassi, C., & Illic, B. M. (2017). The pragmatic turn in studies of linguistic borrowing. *Journal of Pragmatics, 113,* 71-76.

Azam, J., & Mahdieh, N. (2013). Learners use of code switching in the English as a foreign language classroom. *Australian Journal of Basic and Applied Sciences, 7*(1), 186-190.

Bahumaid, S. (2015). Lexical borrowing: The case of English loanwords in Hadhrami Arabic. *International Journal of Language and Linguistics, 2*(6), 13-24.

Bhandari, P. (2020, July 30). *Central tendency: Mean, median and mode*. https://www.scribbr.com/statistics/central-tendency/

Chaudhary, S. (2016). Consumer perception regarding life insurance policies: A factor analytical approach. *Pacific Business Review International, 9*(6), 52-61.

Elliott, A. C., & Woodward, W. A. (2007). *Statistical analysis quick reference guidebook: With SPSS examples* (1st ed.). Sage Publications.

Fareed, M., Samreen, H., & Huma, A. (2016). English language teachers' code-switching in class: ESL learners' perceptions. *Journal of Education & Social Sciences*, *4*(1), 1-11.

Frost, J. (2017). *How to interpret R-squared in regression analysis*. Statistics By Jim. https://statisticsbyjim.com/regression/interpret-r-squared-regression/

Geide-Stevenson, D. (2018). Does English proficiency affect academic performance? *International Review of Economics Education, 28*, 41–48.

Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology Metabolism, 10*(2), 486-489.

Gilakjani, A. P. (2012). A study of factors affecting EFL learners' English pronunciation learning and the strategies for instruction. *International Journal of Humanities and Social Science, 2*(3), 119-128.

Gilakjani, A. P., & Reza, M. A. (2011). A study of factors affecting EFL learners' English listening comprehension and the strategies for improvement. *Journal of Language Teaching and Research*, *2*(5), 977-988.

Haspelmath, M., & Tadmor, U. (2009). *Loanwords in the world's languages: A comparative handbook.* Mouton de Gruyter.

Haugen, E. (1950). The analysis of linguistic borrowing. *Language, 26*(2), 210-231.

Iman, N. B. A. B., Azu, F. A., Khairunnisa, M. K., & Farid, T. M. B. T. A. A. (2015). English-Malay code mixing innovation in Facebook among Malaysian university students. *Researchers World-Journal of Arts, Science & Commerce, 4*(4), 1-10.

Kamisah, A., & Misyana, S. H. (2011). Code-switching and code-mixing of English and Bahasa Malaysia in content-based classrooms: Frequency and attitudes. *The Linguistics Journal, 5*(1), 220-247.

Karim, K., & Nassaji, H. (2013). First language transfer in second language writing: An examination of current research. *Iranian Journal of Language Teaching Research*, *1*(1), 117-134.

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement, 30*(3), 607–610.

Lao, Y. A. D. (2017). First language interferences into English writing skill of the xiith grades students of SMA Negeri 1 Kupang in academic year 2017/2018. *International Journal of Research - Granthaalayah*, *5*(12), 33-44.

Latisha, A. S., & Surina, N. (2013). Languages, code-switching practice and primary functions of Facebook among university students. *Study in English Language Teaching, 1*(1), 187-199.

Lau, S. K., Cheng, X., Tan, K. Y., & Choo, W. L. (2011). Code-mixing of English in the entertainment news of Chinese newspapers in Malaysia. *International Journal of English Linguistics, 1*(1), 3-14.

Lee, Y. L., Ng, Y. J., Chong, S. T., Ariff, M. T., & Nurhanis, S. (2013). Influence of an English lecturer's code-switching practice on students' confidence in the subject. *International Journal of Asian Social Science, 4*(2), 226-233.

Madriñan, M. S. (2014). The use of first language in the second-language classroom: A support for second language acquisition. *Gist Education and Learning Research Journal, 9*, 50-66.

Malah, Z. (2014). *English-Malay loanblends and the indices of adaptability.* International Conference on Languages, Culture and Education, Kuala Lumpur, Malaysia.

https://www.semanticscholar.org/paper/English-Malay-Loanblends-and-the-Indices-of-Malah/e29ed73534b893bae257501d4606eeb1dd45947a

Malaysian Examinations Council. (2017). *Statistics – Official portal Malaysian examinations council.* https://www.mpm.edu.my/en/orang-awam/statistics

Mohammed, R. Z. E. A. L. (2019). Spe*ech accommodation among non-Tlemcenian women in Tlemcen community* [Masters thesis, Aboubekr Belkaid University]. https://www.semanticscholar.org/paper/Speech-Accommodation-among-Non-Tlemcenian-Women-in-Lalimi/bb31bdcaa6cb139de4b94a5d0ba5ff71eba89cd4

Muthusamy, P., Muniandy, R., Kandasmy, S. S., Hussin, O. H., Subramaniam, M., & Farashaiyan, A. (2020). Factors of code-switching among bilingual international students in Malaysia. *International Journal of Higher Education, 9*(4), 332-338.

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing.* United Kingdom: Cambridge University Press.

Nazri, M. L. A. (2013). National language policy and its impacts on second language reading culture. *Journal of International Education and Leadership, 3*(1), 1-11.

New Straits Times. (2019, May 14). Harmonising public and private higher education. https://www.nst.com.my/opinion/columnists/2019/05/488452/harmonising-public-and-private-higher-education#:~:text=Almost%201.3%20million%20Malaysian%20youths,are%20registered%20in%20private%20HLIs.

Ochi, Y. (2009). The role of L1 in facilitating L2 production. *Interpreting and Translation Studies*, *9*, 123-140.

Palacios, N., & Kibler, A. (2016). Oral English language proficiency and reading mastery: The role of home language and school supports. *The Journal of Educational Research, 109*(2), 122-136.

Pillai, S., & Ong, L.T. (2018). English(es) in Malaysia. *Asian Englishes, 20*(2), 147-157.

Rahimi, H. (2014). The effect of method of vocabulary presentation (code-mixing, thematic clustering, and contextualization) on L2 vocabulary recognition and production. *Social and Behavioral Sciences, 98*, 1475 – 1484.

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Llilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 21-33.

Ruth, V. S., & Marcia, A. M. (2018). *Research methods for librarians and educators: Practical applications in formal and informal learning environment*. ABC-CLIO.

Sardar, S. S., Alsamrayee, A. A. M., & Yasin, M. S. M. (2015). Code-switching in daily conversations among Iraqi students in Malaysia. *Arab World English Journal*, *6*(3), 309-319.

Sedgwick, P. (2012). Pearson's correlation coefficient. *BMJ, 345,* e4483. https://doi.org/10.1136/bmj.e4483

Sinha, A., Banarjee, N., Sinha, A., & Shastri, R. K. (2009). Interference of first language in the acquisition of second language. *Journal of Psychology and Counseling*, *1*(7), 117-122.

Smalheiser, N. (2017). *Data literacy: How to make your experiments robust and reproducible*. Academic Press.

Song, J. (2019). Language sociolization and code switching: A case study of a Korean-English bilingual child in a Korean translational family. *International Journal of Bilingualism, 22*(2), 91-106.

Tan, S. I. (2009). Lexical borrowing in Malaysian English: Influences of Malay. *Open Edition Journals, 3,* 11-62.

The Research Advisor. (2006). *Sample size table.* https://www.research-advisors.com/tools/SampleSize.htm

The Star. (2017, January 21). *English proficiency still a big problem for many M'sian grads.* https://www.thestar.com.my/news/nation/2017/01/21/english-proficiency-still-a-big-problem-for-many-msian-grads/

The Star. (2020, January 31). *English to be used to teach Math, Science again, says Dr M.* https://www.thestar.com.my/news/nation/2020/01/31/english-to-be-used-to-teach-math-science-again-says-dr-m

Thirusanku, J., & Melor, M. Y. (2012). The many faces of Malaysian English. *International Scholarly Research Network, 2012, 1-14.*

Thirusanku, J., & Melor, M. Y. (2013). Lexical borrowings by Malaysian ESL teachers. *Asian Social Science*, *9*(13), 185-195.

Thirusanku, J., & Melor, M. Y. (2014). Status of English in Malaysia. *Asian Social Science*, *10*(14), 254-260.

Tsukamoto, M. (2011). Students' perception of teachers' language use in an EFL classroom. *大阪女学院大学紀要,* 8, 143-154.

University of Southern Queensland. (2016). *English Language Proficiency.* https://policy.usq.edu.au/documents/161460PL

Varshney, P. (2020, April 15). *Q-Q plots explained.* https://towardsdatascience.com/q-q-plots-explained-5aa8495426c0

Vollmann, R., & Soon, T. W. (2019). The sociolinguistic status of Malaysian English., 133-150.

Younas, M., Sumrah, A., Khadija, A., Faisal, M., Akhtar, M., Komal, S., & Aneela, U. (2014). Code-switching and code-mixing: A case of EFL teachers affecting L2 learners' learning. *Language in India, 14*, 516-530.

Zulkifli, M. N., Arzilah, N. I., Nuraini, K., Shahrum, A., & Marzuki, M. M. (2011). MUET score and loadings hour: An analysis on the relationship towards academic performance. *Procedia Social and Behavioral Sciences, 18*, 103-109.

APPENDICES

Appendix A

Questionnaire

Please (√) your answers.

Definition of Terms

Code mixing is the mixture of two different languages in terms of words or grammatical features in a particular sentence (Muysken, 2000).

Code switching occurs when a person begins their conversation in one language and shifts to another language halfway, or when a person uses one language but receive replies in a different language (Azam & Mahdieh, 2013).

First language interference is the hindrance of a speaker's first language when learning the second language (Lao, 2017).

Part 1: Demographic Profile (Adapted from Ahmad & Shaima, 2016)

1. Age
    - □ Below 20 years old
    - □ 20 – 21 years old
    - □ 22 – 23 years old
    - □ 24 – 25 years old
    - □ Above 25 years old
2. Gender
    - □ Male
    - □ Female
3. Main medium of education in primary school

□ English

□ Malay

□ Chinese

□ Indian

□ Others: _____ (please specify)

Part 2: Dependent Variable

English Proficiency of University Students in Malaysia

5 – Strongly agree	4 – Agree	3 – Neutral	2 – Disagree	1 – Strongly disagree

No.	Questions	5	4	3	2	1
1.	You fully understand English grammar and are able to use the English language with minimal to no grammatical mistakes.					
2.	You have sufficient vocabulary to help you speak fluent English every day.					
3.	You are able to speak English fluently, accurately and have no problem with pronunciations.					
4.	When listening to English songs or watching English movies, you are able to fully understand them without the help of subtitles.					
5.	When reading English novels, poems etc., you fully understand the context and vocabularies used.					

Part 3: Independent Variable

Code Mixing

1. Do you think code mixing has affected your English proficiency?

   □ Yes

   □ No

2. You subconsciously code mix other language when using English.

   □ Yes

- ☐ No

3. You regularly mix words from other language when using English in your daily conversation.
    - ☐ Strongly agree
    - ☐ Agree
    - ☐ Neutral
    - ☐ Disagree
    - ☐ Strongly disagree
4. You heavily code mix during informal writing (text messages, social media etc.).
    - ☐ Strongly agree
    - ☐ Agree
    - ☐ Neutral
    - ☐ Disagree
    - ☐ Strongly disagree

Code Switching

1. Do you think code switching has affected your English proficiency?
    - ☐ Yes
    - ☐ No
2. You subconsciously code switch from English to another language.
    - ☐ Yes
    - ☐ No
3. You often code switch in your daily conversation.
    - ☐ Strongly agree
    - ☐ Agree
    - ☐ Neutral

- ☐ Disagree
- ☐ Strongly disagree

4. You heavily code switch during informal writing (text messages, social media etc.).
   - ☐ Strongly agree
   - ☐ Agree
   - ☐ Neutral
   - ☐ Disagree
   - ☐ Strongly disagree

First Language Interference (Adapted from Alia et al., 2019)

1. You use bilingual dictionary or direct translations to look for English words that you do not understand.
   - ☐ Yes
   - ☐ No

*2. You directly translate from your First Language to English when writing or speaking.
   - ☐ Strongly agree
   - ☐ Agree
   - ☐ Neutral
   - ☐ Disagree
   - ☐ Strongly disagree

3. You often construct sentences in your First Language first and then translate it to English.
   - ☐ Strongly agree
   - ☐ Agree
   - ☐ Neutral

- ☐ Disagree
- ☐ Strongly disagree

*4. You tend to apply the grammatical rules of First Language in structuring English.

- ☐ Strongly agree
- ☐ Agree
- ☐ Neutral
- ☐ Disagree
- ☐ Strongly disagree

Appendix B

Results of Pilot Tests

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.701	.725	3

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
MEAN_EP	.184	30	.011	.933	30	.058
MEAN_CM	.147	30	.099	.935	30	.067
MEAN_CS	.143	30	.117	.943	30	.112
MEAN_L1	.214	30	.001	.935	30	.066

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.483[a]	.233	.145	.28900

a. Predictors: (Constant), MEAN_L1, MEAN_CM, MEAN_CS

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.660	3	.220	2.636	.071[b]
	Residual	2.172	26	.084		
	Total	2.832	29			

a. Dependent Variable: MEAN_EP

b. Predictors: (Constant), MEAN_L1, MEAN_CM, MEAN_CS

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	1.923	.413		4.661	.000
	MEAN_CM	.424	.194	.565	2.188	.038
	MEAN_CS	-.564	.219	-.688	-2.579	.016
	MEAN_L1	-.049	.117	-.078	-.419	.678

Correlations

		MEAN_EP	MEAN_CM	MEAN_CS	MEAN_L1
MEAN_EP	Pearson Correlation	1	.669**	-.595**	-.174
	Sig. (2-tailed)		.878	.513	.357
	N	30	30	30	30

MEAN_CM	Pearson Correlation	.669**	1	.746**	.504**
	Sig. (2-tailed)	.878		.000	.528
	N	30	30	30	30
MEAN_CS	Pearson Correlation	-.595**	.746**	1	.573**
	Sig. (2-tailed)	.513	.000		.042
	N	30	30	30	30
MEAN_L1	Pearson Correlation	-.174	.504**	.573**	1
	Sig. (2-tailed)	.357	.128	.042	
	N	30	30	30	30

**. Correlation is significant at the 0.05 level (2-tailed).

Appendix C

Reliability Test

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.805	.806	3

Appendix D

Normality Test

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
MeanEP	.153	384	.000	.959	384	.000
MeanCM	.142	384	.000	.962	384	.000
MeanCS	.157	384	.000	.959	384	.000
MeanL1	.163	384	.000	.959	384	.000

a. Lilliefors Significance Correction

Appendix E

Pearson's Correlation Coefficient

Correlations

		MeanEP	MeanCM	MeanCS	MeanL1
MeanEP	Pearson Correlation	1	.583**	.895**	.670**
	Sig. (2-tailed)		.000	.000	.000
	N	384	384	384	384
MeanCM	Pearson Correlation	.583**	1	.575**	.489**
	Sig. (2-tailed)	.000		.000	.000
	N	384	384	384	384
MeanCS	Pearson Correlation	.895**	.575**	1	.679**
	Sig. (2-tailed)	.000	.000		.000
	N	384	384	384	384
MeanL1	Pearson Correlation	.670**	.489**	.679**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	384	384	384	384

**. Correlation is significant at the 0.05 level (2-tailed).

Appendix F

Model Summary of Multiple Linear Regression

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.995[a]	.991	.991	.04451

a. Predictors: (Constant), MeanL1, MeanCM, MeanCS

Appendix G

Analysis of Variance (ANOVA)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	79.532	3	26.511	13382.843	.000[b]
	Residual	.753	380	.002		
	Total	80.285	383			

a. Dependent Variable: MeanEP

b. Predictors: (Constant), MeanL1, MeanCM, MeanCS

Appendix H

Coefficients of Multiple Linear Regression

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	.043	.016		2.637	.009
	MeanCM	.021	.007	.019	3.067	.002
	MeanCS	.980	.007	.994	135.916	.000
	MeanL1	-.016	.008	-.014	-1.978	.049

a. Dependent Variable: MeanEP